



AN ERGONOMICS STUDY OF THE PROCESS OF TAGGING DIGITAL RECORDS USING MACHINE READABLE CONTROLLED VOCABULARIES

Researched and written by Lesley Mackenzie-Robb, MA., BSc (Hons).

Vantaggio Ltd

www.vantaggio-learn.com

©. Vantaggio Ltd 2010

Abstract

This report examines the task of applying a machine-readable vocabulary of terms to digital records to enable their discovery using search engines. The task is analysed using the HTA method (Shepherd, 1998) with the focus on two of the most cognitively demanding sub-tasks involved in the process: term search and term selection. The limitations of the human information processing system are shown to have a potential negative influence on the former, while the decision-making nature of term selection is linked to self-esteem. These factors are correlated to psychosocial factors including the impact of “noise” and “irrelevant speech” in a busy office environment, “time pressure” and “worker control”. These, together with the implications of individual personality differences between operators suggest that a new model for this task is required: this would involve a more controlled environment, individual control over work parameters and the use of an automated system for tagging. It is acknowledged that there is no universally applicable ideal solution. The new model would require evaluation, and measures of evaluation are suggested. Further research is needed in this area to understand the role of individual differences and psychosocial factors in particular.

Introduction

Google estimates that there are more than a trillion distinct URLs on the World Wide Web (Google, 2008). In a successful knowledge economy it is vital that people are able to find the information they need, when they need it. This introduces a two-faceted problem: how to make information discoverable on the web, and how to accommodate for variation in language and terminology?

A widely adopted solution is the use of machine-readable, multi-lingual controlled vocabularies to categorise and describe digital records. A keyword search using an engine such as Google will return records “tagged” with matching and related vocabulary terms. Vocabularies can also be used to aide user navigation and choice in browsing hierarchies.

The process of “tagging” is the application of vocabulary terms to records. The objective is to accurately categorise and describe the record so that it can be discovered by people using search engines or browsing hierarchies. It is usually performed manually, although automatic tagging systems are also in use. There are four problems associated with tagging: under-description, over-description, errors and inconsistencies. The nature of tagging is well understood: what is less well-defined is the extent to which the task is best fitted to the human operator.

Ergonomics focuses on the interaction of human abilities with task and environmental factors (Smith, 1999), and psychosocial factors (Carayon & Lim, 1999). This report studies a tagging task performed in a busy open plan administrative office, based on observation and verbal protocols (Matthews, Davies, Westerman & Stammers, 2000). In this Local Government organisation, most taggers are classed as “information assistants” within Information Systems departments. They have no specialist training in the use of vocabularies but have good computing skills. The task is analysed using the Hierarchical Task Analysis (HTA) method (Shepherd, 1998), and placed within a psychological context. The report concentrates on the two most cognitively demanding elements of the task: a detailed HTA can be found in Appendix 2. The psychosocial factors are evaluated and inter-relationships highlighted. Based on the literature, it can be assumed that there will be a number of influencing factors: individual differences (Ward & Allport, 1997; Josephs, Larrick, Steele & Nisbett, 1992), work environment (Robertson, Huang, O’Neill & Schleifer, 2008) and cognitive load (Sweller, 1994; Paas, Renkl & Sweller, 2003). Options for solutions more suited to the human operator are discussed, and directions for future research identified.

Finding term descriptors

The HTA uses the example digital record titled “ADHD Support Group” (Appendix 1). Records consist of titles and textual descriptions. The aim is to select appropriate descriptive terms from the vocabulary and apply them. This must be balanced with a perception of language used by information-seekers. The operator works in a busy open-plan office, and has other duties in addition to tagging. She is also working to a strict time deadline of tagging a set number of records per day.

The first step involves reviewing the record title and text, scanning for keywords that can be termed “descriptors”. Simultaneously, the operator scans her own knowledge for any related words or synonyms. Thus, there is an element of personal expression involved. In this way, the operator uses the text and her own knowledge to assemble a list of possible descriptors.

In psychological terms, this involves text comprehension, which *may* not compromise the limits of the human information processing system despite the involvement of several parallel processes (Kintsch & van Dijk, 1978). The record title acts as a semantic primer (Eysenck & Keane, 2000), which facilitates the activation of relevant schema in Long-Term Memory (LTM), and which places the text in context (Eyrolle, Verbel & Lemarie, 2008). Eyrolle *et al.* (2008) argue that misleading or incomplete titles can result in processing difficulties. In scanning text for keywords, attention is

focused and goal-driven, which requires effort (Eysenck *et al.* 2000). But attention is not necessarily focused because the operator is also engaged in recalling relevant words from LMT, which implies divided attention. According to Matthews, Davies, Westerman and Stammers (2000) this can lead to attention overload. This part of the tagging task can be seen as an example of divided attention where the operator attempts to maintain several hypotheses in Working Memory (WM) while continuing to scan for other relevant stimuli (Wickens & Hollands, 2000). The limits of WM are such that only around four propositions can be maintained simultaneously, and has a duration of some 20 seconds (Wickens *et al.* 2000). In this case, the record title alone results in two propositions: “ADHD” and “Support”. In the context of Sweller’s (1994) Cognitive Load Theory, the high cognitive demands of this task step increase the risk of poor comprehension (Wickens *et al.* 2000). There is also the possibility of negative performance impact according to the Bottleneck paradigm (Liu, Doong, Hsu, Huang & Jeng, 2009).

Decision-making

Having produced a set of potential descriptor terms, the second sub-goal is to determine if they exist in the vocabulary and if they have the appropriate attributes. The operator uses an interactive, searchable version of the vocabulary to perform this. The operator has to decide if the term location in the vocabulary and its hierarchical associations are appropriate to the record to be tagged. Then, each term’s unique metadata record is examined for information about mapping and term relationships which may influence term selection.

This step principally involves decision-making, and also carries a high cognitive load, with performance potentially limited by an individual’s ability to select one option from several (Ward & Allport, 1997). Studies show a link between decision-making and the psychosocial factor self-esteem (Josephs, Larrick, Steele & Nisbett, 1992). Eysenck and Keane (2000) suggest that “Decision making undoubtedly depends in part on individual differences in personality,” (p 486). Josephs *et al.* (1992) demonstrate this in their study of low and high self-esteem participants. Where low self-esteem participants anticipated feedback on their decisions, they tended to take less risky decisions. In the tagging task, the operator knows that feedback will be given. Recall also that there is an element of personal expression in the task. In the record example, the “safe” decision is to select only “ADHD”, while a more “risky” approach is to select this term in addition to terms associated with disorder symptoms such as “behavioural problems” and “sleeplessness”. It is more risky because these terms can refer to information other than ADHD: in this case, records might be discovered which are irrelevant to the seeker’s enquiry. It has also been suggested that decision-making or planning difficulty increases with the number of choices, particularly where they are perceived as being of equal preference (Ward *et al.* 1997). These two issues – self-esteem and choice preference difficulties – could both be influencing factors in the problem of over and under-descriptions of records. Time pressure can also be a factor, where biases and the use of familiar decision routines can result in error (Wickens & Hollands, 2000).

Once selected, the application of terms to records is a simple computer based process.

Work environment and Psychosocial Factors

There are three key factors which influence performance of the tagging task: noise and irrelevant speech, time pressure and job control.

This particular tagging task takes place in a busy open plan work environment. Noise and irrelevant speech can have a direct and negative impact on WM: complex tasks involving greater cognitive load, as has been shown with the tagging task, are vulnerable to interference and degradation at higher stress arousal levels (Wickens & Hollands, 2000) such as noisy work environments (Radwin & Weneck, 2004). Matthews, Davies, Westerman and Stammers (2000) caution that there is considerable variation amongst studies of noise impact on performance, and that this is probably due to individual differences. They do, however, conclude that "...conversation, in the form of irrelevant speech, is likely to be a much more important contributor to performance decrements," (Matthews *et al*, 2000: p 190), and that studies have consistently shown this (Jones, 1994).

The second factor concerns time pressure. Time constraints for completion of a task can act as a stressor and lead to accelerated work patterns, which can result in more errors being made (Wickens & Hollands, 2000).

The impact of the third factor, job control on performance, is well represented in the literature. The ability to make decisions and take control over the nature of work tasks has been shown to have beneficial implications for health (Dugdill, 2000) and performance (Robertson, Huang, O'Neill & Schliefer, 2008). Robertson *et al.*'s (2008) study of administration workers showed that giving workers control over their work environment also led to a reduction in business process times, although their study's methodology makes generalisation problematic.

These factors could all negatively impact performance and account for the issues associated with tagging.

Individual differences

Individual differences influence performance (Matthews, Davies, Westerman & Stammers, 2000). The first step, for instance, involves the operator's own knowledge. Levels of knowledge and readiness of access in LTM will vary between individuals, and could lead to tagging inconsistencies. Individual self-esteem has already been noted as a marker for individual difference. In the work environment, Wickens and Hollands (2000) note that "...human performance response to stressors appears to be inconsistent and unpredictable.." (p 482), and is influenced by cognitive, skill and personality variables. Ward and Allport (1997), in their study of planning and problem-solving, note that individual differences need to be factored into such studies, while Kintsch and van Dijk (1978) accept that individual difference factors will influence their proposed model of text comprehension. Given the inter-relationship between cognitive and psychosocial factors, and individual differences, to what extent can a single specified process for tagging provide a good fit for all individuals tasked with this procedure?

Solutions and their evaluation

Based on this analysis, the following solutions are offered. First, tagging should be undertaken as a task separate from other tasks and performed in a quiet workspace with regular break intervals. Operators should, in addition, be given control over when, where and how they perform the task.

These two measures directly address the identified psychosocial factors. Thirdly, tagging should be performed using an automated tagging system. This would automate all of step one, and most of step two. The role of the tagger becomes one of quality assurance checker. Matthews, Davies, Westerman and Stammers (2000) are ambivalent over the efficacy of automated systems: assumptions that these will reduce human error and mental workload may be groundless. Whilst this is debateable, it is the case that automated systems will shift responsibility (to some extent) from the operator to the system designer and the vocabulary producer, while introducing a measure of consistency in tagging.

Evaluation of this approach is straightforward. Volumes of records tagged can be compared with the manual method. Consistency of tagging can be compared between taggers. Job satisfaction measures will determine operator motivation and attitudes. Search logs of before and after automation can be compared.

Conclusions

Tagging is a complex, cognitively demanding process, the performance of which is influenced by psychosocial factors such as irrelevant speech and self-esteem. It is unlikely that a perfect, universal solution can be applied due to factors such as individual differences. A new model proposes performance in a quiet environment, using an automated system, and giving the operator more control over their work parameters. Research is needed to determine the value of this approach, and to understand the influence of individual differences and psychosocial factors in more detail.

References

Carayon, P. and Lim, S. (1999). Psychosocial work factors. In K. Karwowski and W. Marras (Eds.), *The Occupational ergonomics handbook*. Florida: CRC Press.

Dugdill, L. (2000). Developing a holistic understanding of workplace health: the case of bank workers. *Ergonomics*, **43**, (10), pp 1738 – 1749

Eyrolle, H., Virbel, J. and Lemarie, J. (2008). Impact of incomplete correspondence between document titles and texts on users' representations: a cognitive and linguistic analysis based on 25 technical documents. *Applied Ergonomics*, 39, pp 241 – 246

Eysenck, M. and Keane, M. (2000). *Cognitive Psychology: a student's handbook*. 4th Edn. Hove: Psychology Press.

Google (2008). We knew the web was big...[Online] Google.
<http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html> Accessed 7 March 2010.

Jones, D. (1994). Disruption of memory for lip-read lists by irrelevant speech: further support for the Changing State hypothesis. *The Quarterly Journal of Experimental Psychology*, **47A**, (1), pp 143 – 160

- Josephs, R., Larrick, R., Steele, C. and Nisbett, R. (1992). Protecting the self from the negative consequences of risky decisions. *Journal of Personality and Social Psychology*, **62**, (1), pp 26 – 37
- Kintsch, W. and van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review*, **85**, (5), pp 363 – 394
- Liu, C., Doong, J., Hsu, C., Huang, W. and Jeng, M. (2009). Evidence for the selective attention mechanism and dual-task interference. *Applied Ergonomics*, **40**, pp 341 – 347
- Matthews, G., Davies, D., Westerman, S. and Stammers, R. (2000). *Human performance: cognition, stress and individual differences*. Hove: Psychology Press.
- Paas, F., Renkl, A. and Sweller, J. (2003). Cognitive load theory and instructional design: recent developments. *Educational Psychologist*, **38**, (1), pp 1 – 4
- Radwin, R. and Weneck, R. (2004). Environment. In (Eds unknown) *Kodak's ergonomic design for people at work*. Chichester: Eastman Kodak Company.
- Robertson, M., Huang, Y., O'Neill, M. and Schleifer, L. (2008). Flexible workspace design and ergonomics training: impacts on the psychosocial work environment, musculoskeletal health, and work effectiveness among knowledge workers. *Applied Ergonomics*, **39**, pp 482 – 494
- Shepherd, A. (1998). HTA as a framework for task analysis. *Ergonomics*, **41**, (11), pp 1537 – 1552.
- Smith, P. (1999). Cognitive factors. In K. Karwowski and W. Marras (Eds.), *The Occupational ergonomics handbook*. Florida: CRC Press.
- Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and instruction*, **4**, pp 295 – 312
- Ward, G. and Allport, A. (1997). Planning and problem-solving using the five-disc Tower of London Task. *The Quarterly Journal of Experimental Psychology*, **50A**, (1), pp 49 – 78
- Wickens, C. and Hollands, J. (2000). *Engineering Psychology and Human Performance*. 3rd Edn. New Jersey: Prentice Hall.